DOCUMENT RESUME

ED 382 005 FL 022 898

AUTHOR Griffee, Dale T.
TITLE A Longitudinal Study of Student Feedback: Self-Assessment, Course Evaluation and Teacher Evaluation.
PUB DATE [95]
NOTE 16p.
PUB TYPE Reports - Research/Technical (143) -- Tests/Evaluation Instruments (160)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS College Second Language Programs; College Students; Conversational Language Courses; *Course Evaluation; *English (Second Language); *Feedback; Foreign Countries; Higher Education; Instructional Improvement; Longitudinal Studies; Second Language Instruction; *Self Evaluation (Individuals); *Student Attitudes; *Student Evaluation of Teacher Performance; Surveys
IDENTIFIERS *Japan

ABSTRACT

An exploratory, longitudinal study was undertaken to investigate how student self-assessment, course evaluation, and student evaluation of teaching can be incorporated into university second language instruction. Three issues were addressed: whether students evaluate themselves differently in first, second, and third year classes in the same university department; how many and what kinds of suggestions students offer for course improvement; and the values that guide students' evaluation of teachers. Subjects were 44 students in one section each of first-, second-, and third year English classes in a Japanese university. All were administered a nine-question survey (appended) concerning their own performance in class, suggestions for course change, and the teacher's performance. Results indicate that, in general, there was little difference between students' self-evaluations at the three instructional levels; students tended to gain confidence and skills as the year progressed. Most students made no suggestions for course change, offering only slightly more as the course progressed. When opinions were expressed, they tended to favor more conversation and music. Comments also suggested that students valued traditional teaching behavior and communicative techniques. A brief bibliography is included. (MSE)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 382 005

A LONGITUDINAL STUDY OF STUDENT FEEDBACK: SELF-ASSESSMENT, COURSE EVALUATION AND TEACHER EVALUATION

Dale T. Griffee

Seigakuin University

Continuing interest in communicative language teaching has raised the issue of negotiation in the classroom, and student evaluation feedback is one form that classroom negotiation might take. Although interest in evaluation feedback has come primarily from North American universities in content areas, recently, there has been increased interest among foreign language (FL) teachers (Maurice, 1992; Yonesaka, 1993). Since FL teaching is by its nature a cross-cultural endeavor with built-in difficulties and possible cultural misunderstanding (see Krasnick, 1991), there is a special need to investigate student feedback as one way to address teacher/student cultural misperceptions especially given the unequal power distribution between teacher and student.

At present, there are typically two avenues of evaluation. One is evaluation of the teacher by the institution (Maurice, 1992, p. 34), which this paper does not address, and the second is evaluation of the student by the teacher. Teacher evaluation of students often takes the form of testing, grades, oral error correction, or written comments on student work by the teacher (Hyland, 1990). What is often missing is student self-assessment, student evaluation of the course content, and student evaluation of the teacher. With no student voice, communication and negotiation are not possible.

The purpose of this exploratory, longitudinal study is to address the question of how student evaluation of self, course, and teacher can be included in the classroom. Three research questions are raised:

1. Is there any difference in student self assessment between first year, second year, and third year classes in the same university department?
2. In terms of class evaluation, how many and what kinds of suggestions do students offer?
3. What values appear to guide students' evaluation of teachers?

Method

Subjects Three general conversation (Eikaiwa) classes were selected from Seigakuin University in the Humanities Department, all classes were in the Euro-American culture division. A first-year, a second-year, and a third-year class were selected with class enrollments of 15, 19, and

FL022898

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.
Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Dale T. Griffee
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

14 (see Table 1). Only one first year class was available, but two sections of the second and third year class were available for this experiment. A single section of the second and a third year class were selected by coin toss. All students were Japanese nationals from 19 to 22 years of age. All classes were on the standard Japanese university 90 minutes, once a week plan. Classes began in April and ended in the first week of July. Classes will resume in September and finish in February 1995.

TABLE 1

Class profile (n = 48)

Class	Enroll.	M	F	Meeting time	Meetings	Absences
1st year	15	8	7	Thursday, 1:30	11	28
2nd year	19	10	9	Monday, 12:50	11	31
3rd year	14	8	6	Tuesday, 8:40	12	29

As can be seen from Table 1, the classes are roughly similar in terms of enrollment, gender, number of meetings, and total number of absences. One difference is that the third year class meets in the morning whereas the other classes meet in the afternoon.

Materials

The feedback instrument (see Appendix 1) is a nine question instrument divided into three parts: student evaluation, class evaluation, and teacher evaluation. Questions one and two were kept vague because asking students about specific class activities raises the problem of metalanguage, the language needed to discuss the activity as opposed to the language used in the exercise. Questions three and seven are scales put in a grade form familiar to all students. These grades were converted to a scale of one to ten. Questions one and two were not delt with in this study.

Procedures

The feedback instrument was administered to each class on the last class day of April, May, and June in 1994. The last twenty minutes of the class were given to this task. Students filled out the forms, put them on the teacher's desk, and left the class with no discussion. The number of forms reported for each month represent the number of students in class on those

dates. Forms were collected and analyzed later. Results of the April survey were reported to all classes, but subsequent results were not reported.

Analysis

Questions 3 (What grade would you give yourself this month?) and 7 (Give the teacher a grade for the month.) were analyzed using the Pearson product-moment correlation coefficient. The three classes were combined and analyzed by month. All correlation coefficients and descriptive statistics were calculated on a Macintosh using SYSTAT 5 for the Macintosh v 5.1.

RESULTS

Of the combined class enrollment of 48, 44 feedback forms were collected in April, 37 in May, and 38 in June (see Table 2). Students gave themselves a grade to cover the one month period of the feedback. This grade was converted to a ten point scale and averaged (see Table 2).

TABLE 2

Student self evaluation grades and student teacher evaluation grades

	April			May			June		
Year	N	self eval	tch eval	N	self eval	tch eval	N	self eval	tch eval
1st	14	5.92	9.35	14	6.78	9.21	11	7.90	9.54
2nd	19	4.84	7.78	16	6.81	8.68	16	7.06	8.87
3rd	11	7.81	8.90	7	8.28	9.85	11	6.45	9.36

Note. Self evaluation and teacher evaluation are marked on a scale of one to ten. One equals a D grade, two equals a C- grade, three equals a C grade up to ten for an A+ grade.

There is a general tendency for students to evaluate themselves more highly as the semester progresses from April to June (see Fig. 1). The one exception is the third year class in June in which self evaluation took an unexpected drop.

Figure 1. Line chart of average of student self evaluation by month

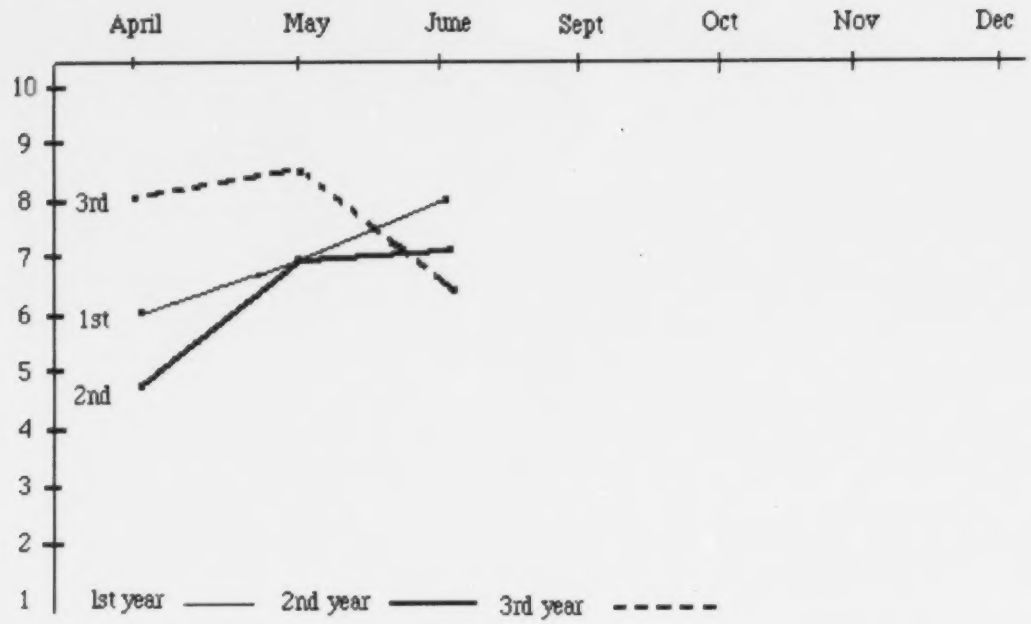

Table 3 shows the number of suggestions converted into percent for comparison. In the first year class in April, 28% of the class made suggestion of some kind, in May 42% made suggestions and in June 45% made suggestions. In the second year class, about half of the class (52%) made suggestions in April, but by June only a quarter of the class did. The third year class was even more erratic starting in April with about one third of the class offering suggestions, then dropping to 14% in May and ending with 27% in June. No clear pattern is discernible in this three month period.

TABLE 3

Number and percent of Class Suggestions

Year	April	May	June
1st			
No suggestions	10	8	6
Suggestions	4	6	5
percent	28	42	45

2nd			
No suggestions	9	14	12
Suggestions	10	2	4
percent	52	12	25
3rd			
No suggestions	7	6	8
Suggestions	4	1	3
percent	36	14	27

Note. Percent is percent of the class who made suggestions.

All suggestions given in answer to question 6 (What are your suggestions for this class?) are listed in Table 4. Two areas seem to stand out, music which received eight mentions and speaking which received seven mentions. At least some portion of all three classes express a desire to engage in more conversation which goes against the accepted wisdom that Japanese students do not wish to engage in conversation in an L2.

TABLE 4

Class Suggestions (Answers to Question 6)

Year	April	May	June
1st	sit in a circle	no textbook	be more friendly
	talk with everybody	play games	listen to music
	have more fun	speak Eng. not Jap.	go outdoors
	have a party	not be absent	don't use tape
		sing songs	more conversation
		(unintelligible)	
2nd	more conversation	listen to music	change working groups
	sit in circle & talk	more speaking	practice speaking
	do role plays		watch movies

	shorter class		listen to tape
	listen to songs		
	listen to music (2)		
	new textbook		
	play games		
	listen to tapes		
3rd	talk more	change rooms	shorter class (2)
	shorter class		more discussion
	listen to music (2)		

Question 7 (What grade would you give to the teacher this month?) is listed by month in Table 5 which also includes the number of comments judged to be negative. These negative comments are listed in appendix 2. High evaluation was operationalized as a grade of A+ (10 points), A (9 points) or A- (8 points). Low evaluation is 7 points or below. The first and the third year classes are similar with infrequent low teacher evaluation (8% and 10%) and high teacher evaluation (92% and 86%). The second year class is more liberal in their criticism of the teacher with low evaluations 27% of the time and high evaluations of 72% of the time.

TABLE 5

Student Evaluation of the Teacher

Year	April	May	June	Total
1st				
no comment	-	-	-	-
negative comments	-	2	-	2
high evaluation	13	11	11	35
low evaluation	1	2	-	3
2nd				
no comment	1	2	1	4
negative comments	5	1	2	8

high evaluation	11	12	14	37
low evaluation	8	4	2	14
3rd				
no comment	1	-	1	2
negative comments	2	-	1	3
high evaluation	9	7	9	25
low evolution	2	-	1	3

Questions 3 (What grade would you give yourself this month?) and 7 (Give the teacher a grade for the month.) were converted from letter to equivalent numbers and analyzed using the Pearson product-moment correlation coefficient. The descriptive statistics and the correlation coefficients are given in Table 6. Both student self evaluation and student teacher evaluation generally improved from April to June. This can be seen in the minimum scores and the mean scores. There is a correlation between ratings students give themselves and the ratings they give the teacher, but it not a high correlation.

TABLE 6

Descriptive Statistics for Answers to Questions 3 and 7 (grade for student and teacher)

	April (n = 44)		May (n = 37)		June (n = 38)	
	Self eval	Teach eval	Self eval	Teach eval	Self eval	Teach eval
min.	1	2	4	6	3	6
max.	10	10	9	10	10	10
range	9	8	5	4	7	4
M	5.96	8.52	7.00	9.11	7.16	9.18
S.D.	2.45	1.95	1.41	1.15	2.10	1.04
skew.	-0.33	-1.45	-0.06	-1.21	-0.69	-1.70
r		0.444*		0.359*		0.359*

Note. min = minimum, max = maximum, M = mean, skew = skewness, r = correlation.

*$p < .05$

DISCUSSION

Research question one was: Is there any difference in student self assessment between 1st year, second year, and third year classes in the same university department? A general answer to this question is no. All classes tended to rate themselves lower at the beginning of the school year and higher as the semester progressed. There are two possible explanations for this. One is that, as the semester progressed, students gained confidence in their ability to perform and also gained familiarity with using the feedback form. Another explanation is that the rise in self evaluations is a product of teacher intervention. After the April feedback forms were received, the teacher carefully examined all forms for all classes and in many cases took remedial action such as having a class discussion of the results. Another example of teacher intervention as a result of the feedback forms is with the second year students. These students often work in groups and one group complained that the teacher did not visit their group often enough. The teacher visited this group in May and these visits by the teacher were noted by these students in the May feedback evaluation. The third year students evaluated themselves the highest of the three groups in April and the lowest of the three groups in June. Because the feedback forms were not, excepting for April, evaluated on a monthly basis, an opportunity to discuss this drop in the June self evaluation was missed, which is unfortunate because the reason for the drop might have revealed something interesting about the class or the lives of the students. This points to the need for an evaluation procedure at the end of every month.

The second research question was: In terms of class evaluation, how many and what kinds of suggestions do students offer? First, we will examine the "how many" issue and then examine the "what kinds" issue. One interesting finding of this study is that the majority of students in all classes do not make any suggestions at all. The silent majority is a reality. For example, the first year students begin with only 28 percent of the class offering suggestions in April and then gradually increase to 45 percent by June. One explanation for the first year students is that they took a "wait and see" attitude in what was for them their first semester of their first year of university English. The second and third year students both begin with relatively high numbers of suggestions, then drop the next month, and then have a slight increase the last month.

In terms of what kinds of suggestions students make, the frequent number of requests for more conversation and/or discussion is notable as is the request for more songs and music.

This suggests that the needs of a certain portion of the students are not being met which raises the possibility that each class in this study is composed of subgroups of students who are at different levels and perhaps have different interests. In other words, instead of their being, for example, one third year class, there may be two or three distinct groups in the third year class each of which has its own needs and goals. How to identify these subgroups within the larger groups and the creation of appropriate instructional materials for them is beyond the scope of this study, yet clearly needs to be investigated.

In addition to what kinds of suggestions students make, an important issue is how a teacher can deal with student suggestions and criticisms, an issue to be taken up in the conclusion. Here, we will discuss students' negative comments. The first year students offered few negative comments and very high evaluation of the teacher. The second year students offered the most negative comments, but still showed a clear trend from initially low evaluation of the teacher to final high evaluation. The third year class offered fewer negative comments than the second year students, but followed the same tendency. These classes might be characterized respectively as sweet-and-naive (the first year class), angry and disillusioned (the second year class) and, we-see-what-the-situation-is but we-are-not-going-to-rock-the-boat (the third year class). Thirteen comments were judged to be negative (see Appendix 2). Four comments were repeated more than once. These were 1) crying, 2) speaking too fast, 3) the teacher not talking to certain groups, and 4) being absent from class. All of these are self evident except number one, crying. This refers to the teacher simulating crying when students would not ask questions. Although most of the students understood the joke, two students sitting in the back apparently were offended by the sight and sound of their teacher crying which they did not consider appropriate classroom teacher behaviour.

The third research question was: What values appear to guide student evaluation of the teacher? While it is not easy to identify values, the negative comments suggest that students value traditional teacher behavior, being communicative and being present.

Looking at the positive comments, we can say at least three things. First, female students tended to rate the teacher slightly higher than the male students. Of ninety high grade evaluations given to the teacher, females give 51 high grades as compared to the males who gave 43 high evaluations. Second, students tended to rate the teacher somewhat as they rated themselves. In the correlation study run between student self evaluation and student evaluation

of the teacher, the aim was to determine if the evaluation of the teacher was, in fact, a reflection of the student's evaluation of themselves. The correlation figures indicate that this might be the case. It is almost never the case that students evaluated themselves higher than they evaluated the teacher. If a student gave herself a high grade, she tended to give the teacher a high grade and if she gave herself a low grade she also tended to give the teacher a low grade. Third, the ninety-four comments that were made by students who assigned the teacher a high evaluation were divided into eight categories in an attempt to see more clearly what values were guiding the students. These categories were the teacher: 1) was competent, 2) worked hard, 3) was funny or interesting, 4) was supportive or kind, 5) created a good class atmosphere, 6) received a good rating with no reason for the rating being supplied, 7) had good class activities and, 8) had personal qualities the student liked.

Two raters reached a 74.42 percent agreement matching student comments with these eight categories. The major source of disagreement between the raters were categories three and five. Rater one assigned a total of 29 comments to these two categories and rater two assigned 30 comments to these two categories. The difference was that rater one assigned 27 comments to category three and two comments to category five while rater two assigned 15 comments to category three and 15 comments to category five. The majority of comments were assigned by both raters to categories three, five, and six. Category six is a general category which give the teacher a high evaluation with no indication of why. Typical comments in this category are, "You are wonderful" which is gratifying for a teacher to hear, but gives no indication why the student thinks this is the case. Therefore, category six was not considered. The remaining categories with the highest total number of comments are categories three (that the teacher is interesting or funny), five (class atmosphere), and one (teacher judged to be competent) which was a distant third. Rater one assigned eight comments to category one and rater two assigned ten comments. In all other categories, both raters assigned 10 or less of comments. Since category six is factored out because it contains no reasons and thus no values, the predominant values as revealed in this study are that Japanese students value a teacher who is primarily funny or interesting and secondarily a teacher who is helpful. This conclusion corresponds with Durham and Ryan (1991, p. 79) who asked Australian and Japanese students to describe and define their idea of a good teacher and found that the Japanese students wish to be entertained rather than informed. Ryan conducted his study in Japan so it may be assumed that his

Japanese students were reporting on their Japanese teachers. My results would indicate that whether a teacher is Japanese or non-Japanese, students prefer a teacher who is entertaining over a teacher who supplies information.

CONCLUSION

The purpose of this study was to investigate how students evaluated themselves, their course, and their teacher. It was also to determine how a dialogue between students and teacher could be initiated on the same topics. It must be remembered that only three data collection periods have elapsed with at least another three to occur in October, November, and December 1994. This study to date suggests that by using a feedback instrument once a month rather than at the end of a course coupled with the use of student names on the feedback instrument, student/teacher negotiation can be promoted. Without student names on the feedback instrument, remedial action by the teacher could be general but not specific, thus possibly reducing its effectiveness. On the other hand, conventional teacher wisdom suggests that the inclusion of a student's name on a feedback evaluation form might consciously or even unconsciously bias the student especially in the area of teacher evaluation. The high evaluation received by the teacher in this survey might be due to various causes other than teacher excellence including student bias or student worry about grades. Therefore, in future studies, the feedback form will be submitted to the teacher in a blind fashion. First, student names will not be included on the form and second, forms will handed out at the end of the period and the teacher will leave the room leaving the forms to be picked up by a student and returned to the teacher. This later action will remove possible bias resulting from the fact that the student would have to hand the form to the teacher which the teacher might immediately read even though the student did not sign her name.

Another performance feature of the feedback instrument was to profile each of the three classes in the study. The second year class seems different from either the first or third year classes. In fact, the second year students are involved in a pedagogical experiment in which the other two classes are not. At the beginning of the school year, each student in the second year class was required to formulate personal class goals for the year and to report what progress they made toward those goals after each class. The first and third year classes followed a more traditional textbook oriented approach. Perhaps it is the effect of this experiment that caused the difference in the evaluation of the second year class.

Finally, this study highlights the necessity of teacher response to student feedback. It has already been noted that the lack of monthly inspection of the feedback documents caused the teacher to miss the dramatic drop in the third year student self evaluation. As a result, from the beginning of the next school semester in October, monthly tabulation of the averages of student self evaluation and student evaluation of the teacher will occur. In this way, any dramatic rise or fall can be monitored and discussed with the class. In addition, suggestions and negative feedback will also be tabulated for possible teacher action.

REFERENCES

Breen, M. (1991). Understanding the language teacher. In Phillipson, Kellerman, Selinker, Sharwood Smith and Swain (Eds.), Foreign/Second Language Pedagogy Research (pp.213-233). Clevedon: Multilingual Matters.

Donmall, G. (1991). Old problems and new solutions: LA work in GCSE foreign language classrooms. In C. James & P. Garrett (Eds.), Language awareness in the classroom (pp. 107-122). London: Longman.

Durham, M. & Ryan, S. (1991). What kind of teacher do students want? A cross-cultural comparison. Speech Communication Education, 5, 70-99.

Hyland, K. (1990). Providing productive feedback. English Language Teaching Journal, 44, 279-285.

Krasnick, H. (1991). The use of the concept of cultural sensitivity in teacher evaluation: A case study. TESOL Quarterly, 25, 195-200.

Maurice, K. (1992). Using learner evaluations to improve language teaching. JALT Journal 14, 25-51.

Murphy, D. (1993). Evaluating language learning tasks in the classroom. In G. Crookes & S. Gass (Eds.), Tasks in a pedagogical context (pp. 139-161). Clevedon: Multilingual Matters.

Sharp, A. (1990). Staff/student participation in course evaluation: A procedure for improving course design. English Language Teaching Journal 44, 132-137.

Wennerstrom, A. K. & Heiser, P. (1992). ESL student bias in instructional evaluation. TESOL Quarterly, 26, 271-288.

Yonesaka, S. (1993). Creating student evaluations of instructional forms. The Language Teacher 17, (4), 19-24.

APPENDIX 1

Feedback form

Name _____ number _____

Date _____

Student Feedback & Evaluation

Student evaluation

1. What was your strong point in class this month?_____

_____

2. What was your weak point in class this month? _____

_____

3. What grade would you give yourself this month? Circle one.

A+ A A- B+ B B- C+ C C- D

4. Why? _____

_____

Class evaluation

5. What would you like to change in this class?_____

_____

6. What is your suggestion? _____

_____

Teacher evaluation

7. Give the teacher a grade for the month.. Circle one.

A+ A A- B+ B B- C+ C C- D

8. Why? _____

9. What suggestions do you have for the teacher. _____

APPENDIX 2

Teacher Evaluation Negative Feedback Comments

Year	Month	#	comments
1st	April	-	
	May	2	crying (2)
	June	-	
2nd	April	5	I want the teacher to be more active.
			You spoke too rapidly for me.
			I don't understand what to do.
			Your talking is so funny, but you spoke too fast.
			We don't speak with teacher many times.
	May	1	Please come to my group.
	June	2	Because you were absent one class. (2)
3rd	April	2	You were absent from class.
			You were absent one class this month.
	May	-	
	June	1	Because you can't speak Japanese.